

Representation of spectrotemporal sound information in the ascending auditory pathway

Monty A. Escabi^{1,2}, Heather L. Read³

¹ Biomedical Engineering Program

² Department of Electrical and Computer Engineering

³ Department of Psychology

University of Connecticut, Electrical and Computer Engineering, 371 Fairfield Road U-1157, Storrs, CT 06269-2157, USA

Received: 10 July 2003 / Accepted: 10 September 2003 / Published online: ■

Abstract. The representation of sound information in the central nervous system relies on the analysis of time-varying features in communication and other environmental sounds. How are auditory physiologists and theoreticians to choose an appropriate method for characterizing spectral and temporal acoustic feature representations in single neurons and neural populations? A brief survey of currently available scientific methods and their potential usefulness is given, with a focus on the strengths and weaknesses of using white noise analysis techniques for approximating spectrotemporal response fields (STRFs). White noise analysis has been used to foster several conceptual advances in describing neural acoustic feature representation in a variety of species and auditory nuclei. STRFs have been used to quantitatively assess spectral and temporal transformations across mutually connected auditory nuclei, to identify neuronal interactions between spectral and temporal sound dimensions, and to compare linear vs. nonlinear response properties through state-dependent comparisons. We propose that noise analysis techniques used in combination with novel stimulus paradigms and parametric experiment designs will provide powerful means of exploring acoustic feature representations in the central nervous system.

1 Spectrotemporal decomposition by the auditory system

Sound information is transduced into electrical activity by an array of sensory receptors in the mammalian cochlea. Each receptor, or hair cell, is selective for the frequency of sound vibrations and organized systematically in the cochlea from low to high frequency (Liberman 1982). This frequency information is relayed by the bundle of auditory nerve fibers leading to the

central nervous system, where the sound is further parsed and processed into identifiable messages.

Frequency is the principal organizational parameter for sound information in the auditory pathway. The sensory neuropil of nearly every auditory station is organized into contiguous frequency-tuned channels, preserving the frequency selective axis of the cochlea. By direct analogy the position of a spot of light in a visual scene is expressed in the spatial location of neurons in the retinal epithelium, and this representation is preserved in the central visual system (Tusa et al. 1978). However, while the neuronal organization in the retina of visual stimuli follows a nearly point-for-point conformal mapping of the external visual image, the spatial position of sound sources is not systematically mapped in the early auditory system. Instead, acoustic sound vibrations are converted into a time-frequency pattern of neuronal activity on the sensory receptor surface (Fig. 1).

The principal outcome of this early transformation by the cochlea is that sound signals are not explicitly represented in the brain as they are in the external world. Thus, position in the external world is relayed to the brain through frequency-dependent cues, which have to be sorted and decoded at higher levels of processing. This initial decomposition by the peripheral auditory receptor constrains the representation of sounds to a time-dependent frequency activation pattern. As such, the input to the central auditory system is best expressed as an image pattern that depicts the activity of the cochlea. The central auditory system must, in turn, manipulate this neuronal activity pattern to decipher and extract information about the source content and to estimate the position of the sound source in the external environment.

How is this sound information in the cochlear output pattern represented and processed by the brain? In this review, we will discuss how sound information is decomposed in the central auditory system into elementary spectrotemporal features. We will focus on showing how spectrotemporal components in complex

Correspondence to: M. A. Escabi
(e-mail: escabi@engr.uconn.edu)

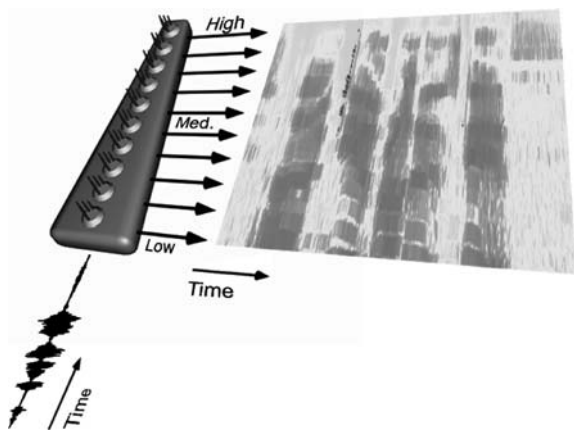


Fig. 1. Spectrotemporal decomposition by the cochlea. An array of frequency-tuned hair cell receptors on the basilar membrane decomposes (unfolded for illustration purposes) the continuous sound vibrations (*black waveform*) into a spectrotemporal activity pattern. Hair cells are aligned from low to high frequencies on the basilar membrane. The spectrotemporal neuronal activity pattern reflects the time-varying energy modulations of the stimulus along each frequency channel from low to high frequency. Illustrated for a segment of human speech

sounds are systematically manipulated and transformed across higher levels of processing.

2 Characterizing auditory receptive fields: first-generation approaches

Stimulus response function

Our understanding of how spectral and temporal information in complex signals is represented in the nervous system has relied on methods to measure the stimulus-response function (SRF) of neurons. The SRF describes the manner in which a sound is encoded by an individual neuron as a train of action potentials or spikes. Unfortunately, the number of available methods to measure neuronal sensitivities is almost as numerous as the number of researchers studying the auditory system – and this presents a significant interpretational challenge. Technically, the types of SRF that can be constructed from neurophysiological data are strictly limited by an experimenter's ability to synthesize novel stimulus paradigms to address a specific question or hypothesis. Apart from this unlimited supply of experimental test stimuli, analysis approaches for measuring neuronal SRFs typically fall into one of three categories: poststimulus averaging, parameter-space approaches, and preevent averaging.

Generally speaking, the SRF of a neuron can be defined as the collection of stimuli, stimulus parameters, or stimulus combinations that elicit specific neuronal response patterns. Ideally, a complete characterization of a single neuron would therefore require that we present every existing sound stimulus in the auditory environment, in every possible combination, and that we measure all the elicited response patterns. That is, not only must we consider all of the possible stimuli that may

potentially elicit responses, but we must in addition identify the action potential response sequences associated with a given stimulus. A practical solution is to probe neural responses with a subset of stimuli to approximate the SRF.

Parametric and poststimulus analysis of SRF

One way to remedy this dimensionality problem is to consider only a few stimulus parameters of interest, such as sound frequency or intensity, from which one can perturb the sound by independently manipulating each parameter. Neuronal response firing rates can then be plotted as a function of each parameter. Figure 2 illustrates such an example for an auditory midbrain neuron from the inferior colliculus that was tested with tone pips of different frequencies and intensities (Fig. 2a). The neuron's frequency tuning curve (FTC) is obtained by plotting the firing rate as a function of frequency and sound pressure level (Fig. 2b). The color scale on the response surface plot designates the neuron's response firing rate as a function of the stimulus parameters (Fig. 2b). For a large subset of the frequencies and intensities tested, the neuron produces no noticeable change in firing rate (blue). The neuron responds over a restricted range of frequencies and intensities and is preferentially activated by sound frequencies around 18 kHz (i.e., the best frequency, or BF). As the sound intensities become louder, the range of frequencies (i.e., the bandwidth, or BW) that activate this neuron broadens. That is, at high sound intensities the neuron becomes less *selective* for the frequency content of the sound.

Parameter-space approaches of this form are extremely useful and intuitive; however, they are usually limited to simple sounds that are easily manipulated by independent parameters. Furthermore, representing the neuronal response as an average firing rate can discard meaningful structure in the temporal response pattern that can potentially convey information about the dynamics of the sensory stimulus.

Poststimulus averaging techniques provide one way of recapturing response timing information in neuronal spike trains. Poststimulus averaging methods, such as the poststimulus time histogram technique, rely on repeated presentation of a fixed stimulus, while the response following the stimulus is collected and averaged using sufficiently small time bins. This repeated averaging over multiple trials removes spike-timing noise that is not systematically related to the stimulus. The primary advantage over the direct parametric approach is that the neuronal response can be studied as a function of time at a chosen temporal resolution. This additional timing information can, in turn, be used to relate the precise timing of action potentials to specific stimulus instances or to measure the reliability and stimulus information content conveyed by single action potentials (de Ruyter van Steveninck et al. 1997). Parameter-space approaches are often combined with poststimulus averaging techniques into hybrid methods that allow exper-

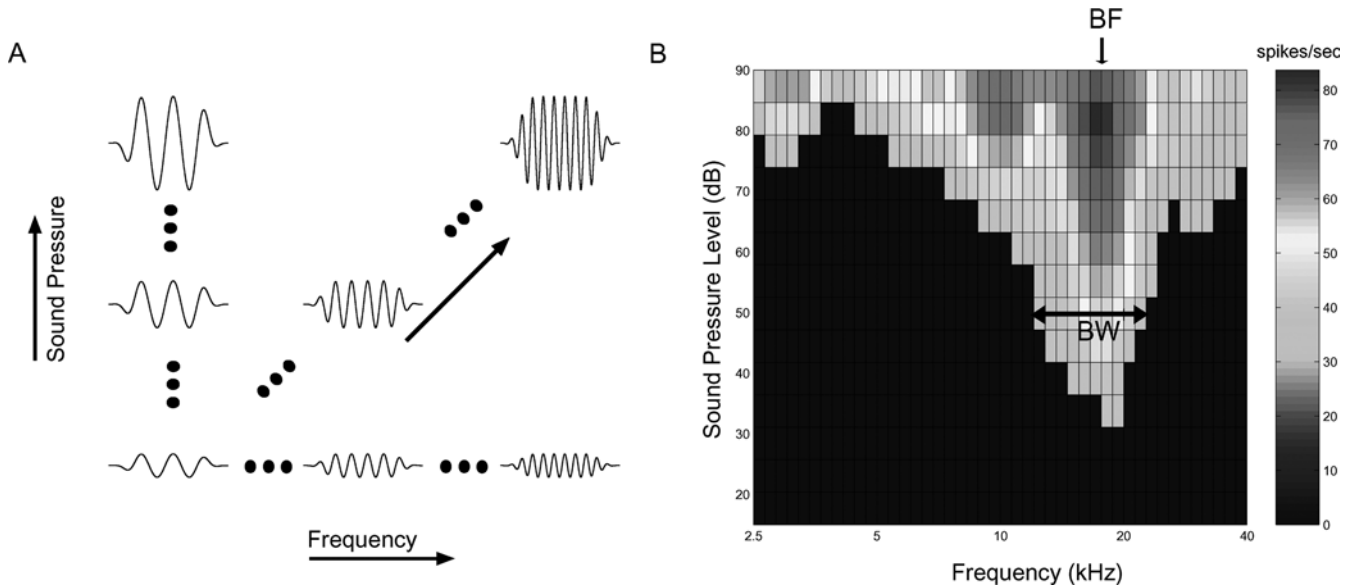


Fig. 2. Conventional frequency tuning curve (FTC) receptive field. Pure tones of varying frequency and intensity (SPL) are presented (a) to the animal. The direction of increasing frequency and sound pressure are designated by arrows. Response firing rates for each condition are tabulated into a two-dimensional matrix (b). The color bar designates the measured neuronal firing rate in spikes/s. This

neuron responds maximally with a best frequency (BF) near 18 kHz. Its frequency response bandwidth (BW) increases with increasing sound pressure, and the neuron is therefore less selective for frequency components at high sound levels. Illustrated for a cat inferior colliculus neuron

imenters to study temporal information processing as a function of a specific stimulus parameter (Langner and Schreiner 1988; Schreiner and Urbas 1988; Rees and Palmer 1989).

Despite their limitations, these approaches have successfully identified the representation of simple acoustic parameters in various auditory nuclei. For instance, frequency selectivity has been probed with tone pips of different frequencies by tabulating the response of neurons as a function of preferred frequency and position in cortical and subcortical areas (as in Fig. 2). These studies have shown topographic cortical and subcortical representations for tone-frequency selectivity (Merzenich and Reid 1974; Merzenich et al. 1975) and for frequency integration bandwidths (Schreiner and Langner 1988; Schreiner and Mendelson 1990; Heil et al. 1992; Read et al. 2001). This frequency-based representation is an integral feature of the entire lemniscal auditory pathway. Similarly, auditory neurons are also selective for temporally varying stimulus components including temporal modulations in signal energy and time-varying frequency transitions (Langner and Schreiner 1988; Schreiner and Urbas 1988; Rees and Palmer 1989). These too can show a systematically distributed organization in both subcortical and cortical stations (Schreiner and Langner 1988; Heil et al. 1992; Mendelson et al. 1993; Read et al. 2001).

Probing the SRF with species-specific sounds

An alternative approach to testing neuronal sensitivities in animal species that rely heavily on communication behavior is to approximate the SRF for animal vocal-

ization by parametrically modifying the structure of the natural vocalization. The most convincing examples showing the importance of spectral and temporal sound structure on neuronal information processing comes from species where a specific audio-vocal behavior guides the functional organization of the sensory representation. Neurons in these model systems are exquisitely tuned to biologically important sounds. Species-specific vocalizations are therefore often recorded onto a computer where they are digitally manipulated (by changing spectral or temporal parameters) and played back to the animal. This approach has been extremely successful for studying the neuronal substrate of prey-velocity detection and spatial navigation in echolocating bats (Suga et al. 1975; Suga and Jen 1976; Fitzpatrick et al. 1993; Casseday et al. 1994; Portfors and Wenstrup 2001). Behavioral specificity in bats is reflected in the magnified representation of combination-sensitive and time-delay-tuned neurons that respond selectively to the frequency content and temporal structure in their echolocation and communication signals. Bats emit high-frequency sonar pulses, which they use to navigate during active echolocation. By comparing the timing delay and spectral energy differences between the emitted sonar pulses and the return echoes, they can create an internal image of their external environment and the location of their prey. Similarly, neuroethologically motivated approaches have also been the principal driving force for defining the neuronal mechanisms underlying song analysis, production, and learning in songbirds. Songbird neurons can be quite sensitive to the sequential order of phrases or temporal conjunctions in a song motif and to temporal stretching or compression of a song sequence

(Brenowitz et al. 1997; Doupe 1997). As such, these manipulations have demonstrated extensive neuronal selectivity to the dynamic spectrum of natural sounds by showing a direct link between sound structure, behavior, and neuronal sensitivity.

There are two problems with using naturalistic stimuli alone to probe neuronal response preferences in most other species. First, not all animals have highly stereotyped vocalizations to draw from, so that selecting the appropriate stimulus or defining appropriate stimulus paradigms is not trivial. Second, as demonstrated in the bat, the use of a limited set of natural sounds with highly biased spectrotemporal frequency distributions to study neuronal sensitivities makes it difficult to predict how the same neurons would respond to other relevant sounds. Cortical neurons in the bat were previously thought to be specialized exclusively for echolocation pulses. However, these neurons actually exhibit dual sensitivity to echolocation and communication sounds (Ohlemiller et al. 1996; Razak et al. 1999), thus serving an important secondary function.

3 Characterizing auditory receptive fields: second-generation approaches

Receptive field (RF) analysis has been progressing in recent years because of the need to study neuronal dynamics of frequency integration with increasingly more complex auditory stimuli including natural sounds (Nelken et al. 1997; deCharms et al. 1998; Klein et al. 2000; Theunissen et al. 2000; Depireux et al. 2001; Schnupp et al. 2001; Sen et al. 2001; Blake and Merzenich 2002; Escabi and Schreiner 2002; Miller et al. 2002; Linden et al. 2003). As stimulus complexity increases, it becomes increasingly more difficult to evaluate neuronal preferences using direct parameter-space approaches because spectral and temporal features are very often not easily represented by simple parameters. Alternative methods are therefore required to evaluate the relationship between sound stimulus and response.

White noise analysis, preevent averaging, and reverse correlation

The combined use of white noise and reverse correlation analysis has been a successful and practical approach to mapping both complex and simple response properties in the auditory system. This method was first introduced into neuroscience to study the temporal filtering properties of single auditory nerve fibers (de Boer and de Jongh 1967). Later, this technique was extended to the analysis of biological systems with multiple inputs (Marmarelis and Naka 1974), such as for the array of photoreceptors in the retina or the collection of hair cells in the cochlea. This led to the development of the spatiotemporal receptive field in the central visual system (Jones and Palmer 1987; DeAngelis et al. 1993) and the spectrotemporal receptive field (both referred to as STRF) in the auditory system (Hermes et al. 1981).

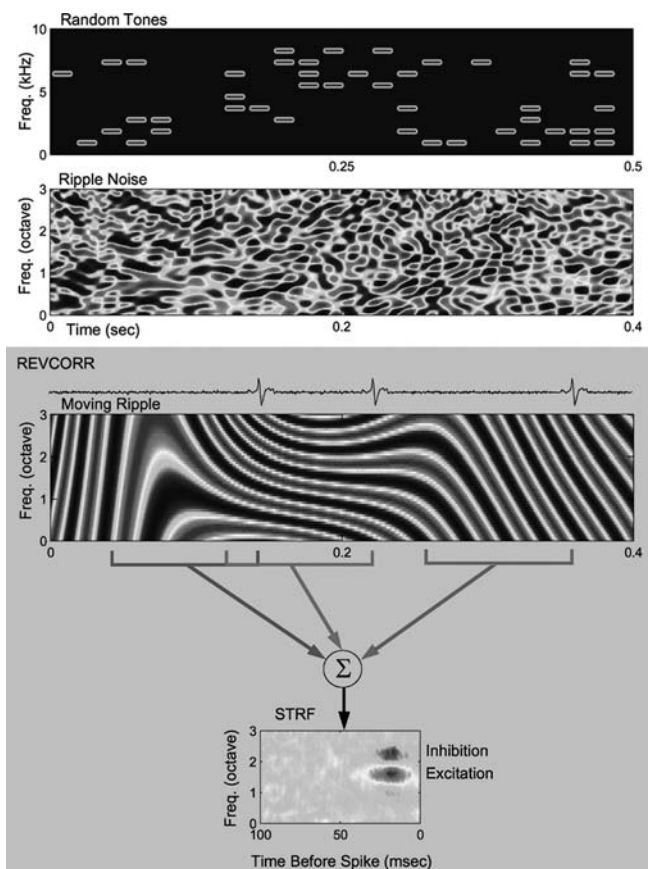


Fig. 3. Estimating the auditory STRF via reverse correlation for various sound ensembles. Example of three sounds (*top*) that are structurally compatible with the STRF measurement procedure: random tone sequence, ripple noise, and dynamic moving ripple. These sounds cover a wide range of spectrotemporal modulations in an unbiased manner, so that the STRF structure is not systematically biased by the stimulus properties. The reverse correlation or REVCORR (*gray box*) is shown for the dynamic moving ripple stimulus. The STRF is estimated by spike-triggered averaging of the sound waveforms just before each action potential. The STRF reflects the preferred stimulus pattern that tended to activate the neuron and is expressed as a time-frequency pattern of neuronal excitation (*red*) and inhibition (*blue*). The example neuron contains a high-energy excitatory response with a frequency of ~ 1.5 octaves and duration of ~ 30 ms and a low-energy inhibitory receptive field subregion of ~ 2.5 octaves

Generating STRFs

Unlike conventional frequency tuning curve receptive field approaches (Fig. 2), which measure spectral integration properties independently of timing, auditory STRFs capture the evolution of the frequency response integration over time. This history of the response can be consolidated into a pictorial representation of the neuron's preferred sound pattern. Generating an STRF first requires a sound that covers a wide range of stimulus possibilities so that the neuron is evenly exposed to a large subset of all possible sound configurations (Fig. 3). The STRF is then obtained through a procedure called reverse correlation or spike-triggered averaging (sometimes also referred to as preevent

averaging; Fig. 3). Reverse correlation is performed by averaging all the stimulus patterns just preceding each action potential (Fig. 3, bottom panel). In this manner, the history of the sound is compared to the history of the neuronal spikes and this reveals the stimulus spectrotemporal configuration that is most likely to activate the neuron. The basic assumption behind this procedure lies in the fact that neuronal responses are not simply random events but are instead the manifestation of sound patterns that efficiently activate the neuron. One can therefore look back in time over a short time interval prior to each spike to get a sense of the sound configuration that causes the neuron to spike. This spectrotemporal averaging therefore reveals the neuron's preferred stimulus, expressed as a time-frequency pattern of sound modulations preceding the onset of action potentials. An example STRF is shown in Fig. 3. Average spiking of the neuron's output is indicated by the color spectrum shown. Red stimulus domains in the STRF indicate that the stimulus tended to be on prior to the initiation of an action potential (at delay 0 ms), whereas blue indicates that the stimulus was, on average, off. The example neuron (Fig. 3) is preferentially activated by a 1.5-octave frequency sound component that is turned on and off sequentially in time. The STRF can alternatively be interpreted as a spectrotemporal filter function, in which case the red receptive field regions correspond to excitatory neuronal inputs that tend to increase the neuron's activity above the mean response level. Blue or off-sound regions are predominantly inhibitory or suppressive, and on-stimuli presented within these domains would decrease the firing rate.

4 Spectrotemporal preferences and transformations

STRF-based techniques provide a means of measuring neuronal response properties at different sensory processing stages that can be characterized and compared quantitatively. Dramatic changes in receptive field properties have been observed in both auditory and visual sensory pathways using conventional SRF and species-specific SRF techniques. The spectrotemporal transformations in the transition from thalamus to cortex have been recently confirmed and extended using STRF analysis. Here we will discuss a series of studies that have examined STRF preferences and the transformation across various auditory nuclei.

Simple features of STRFs

One of the advantages of white noise and reverse correlation techniques is that spectral and temporal response properties can be thoroughly quantified together. Furthermore, the excitatory and inhibitory nature of the response can be approximated. There are several ways to quantify temporal and spectral sensitivities and, as discussed later, their interactions. Simple structural parameters can be measured directly from the

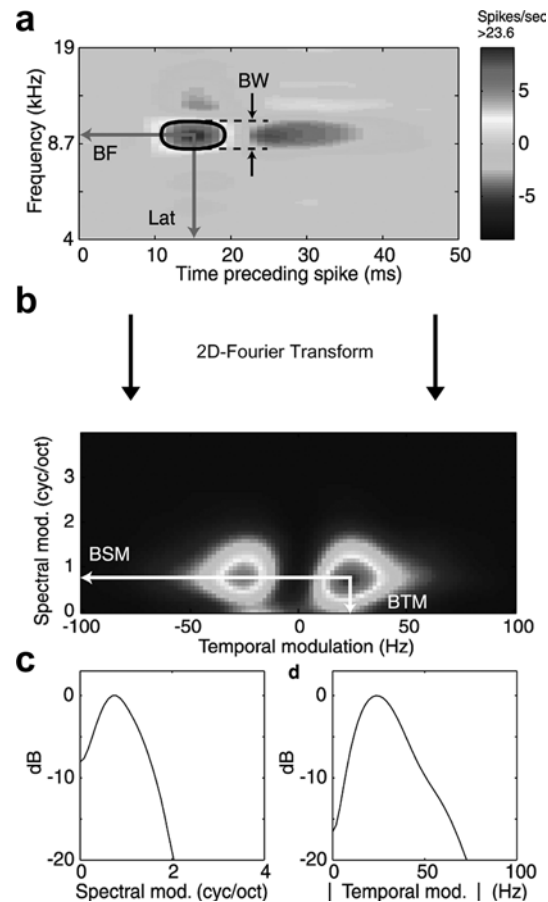


Fig. 4. Relating the STRF to the neuron's spectrotemporal modulation preferences. **a** Simple response parameters such as the response latency (Lat), best frequency (BF), and receptive field bandwidth (BW) are immediately available from the STRF (black and gray arrows). These parameters, however, focus exclusively on the short-latency dominant excitatory receptive field subregion (encircled black region). **b** Other, more intricate modulation response parameters can be obtained by converting the STRF into a frequency domain representation: the ripple transfer function (RTF). Here, a two-dimensional Fourier transform is used to convert the STRF into constituent spectral and temporal parameters. The time axis of the STRF is converted into a temporal modulation frequency axis while the spectral axis is converted into a spectral modulation axis. The location of the dominant energy peak of the RTF depends strongly on the spectrotemporal arrangements and the extent of excitation and inhibition in the STRF. This peak determines the neuron's best temporal modulation (bTM) and best spectral modulation (bSM). The depicted thalamic neuron responds preferentially with a bTM of ≈ 25 Hz and bSM of ≈ 0.75 cycles/octave. Spectral (**c**) and temporal (**d**) cross sections of the RTF show that the neuron selectively filters spectrotemporal modulations around the bTM and bSM

STRF. Two types of temporal properties can be measured directly as the time to maximum excitatory response (i.e., peak latency) or the overall response duration in the STRF (Fig. 4a). Spectral RF attributes are easily obtained by measuring the frequency associated with the maximal mean firing rate (i.e., a neuron's best frequency, BF) and the total range of frequencies within the significant portion of the STRF (i.e., receptive field bandwidth, BW). Conceptually, peak latency, best frequency, and bandwidth are analogous to the same

parameters obtained with more conventional techniques such as the FTC (Fig. 2) and histogram analysis of FTC spike data (Cheung et al. 2001). Although some RF properties appear to be the same between STRFs and FTCs, other response properties are not. First, STRFs often show on and off (excitatory/inhibitory) temporal response patterns extending over tens to hundredths of milliseconds. The duration of the excitatory RF can be measured with single-tone FTC spike data, but inhibition or suppression of spike output cannot be measured without the addition of a masking or priming stimulus (Brosch and Schreiner 1997). Second, although receptive field best frequencies are very similar for STRF and FTC, receptive field bandwidth relationships are often significantly different for the two approaches. Bandwidth typically increases with increasing SPL for pure tones (as in Fig. 2); however, this is not the case for STRFs obtained with broadband noise that have constant bandwidths with increasing SPL (Escabí et al. 2003; Valentine and Eggermont 2003). Such differences indicate that narrow band pure tones and broadband sounds that fully engage the cochlear sensory epithelium are not entirely complementary. As such, the behavior of the neuron's response and its preferences to pure tones is insufficient to explain the neuron's preference to more complex sounds and vice versa. Until the underlying cause of such RF differences is known, it is of some advantage not to abandon conventional approaches altogether.

Extracting spectral and temporal features from the STRF

The simple descriptions of STRF parameters provide a tool for quantifying basic measures of neuronal sensitivity and for relating these to conventional SRF approaches. However, these parameters tell us nothing about more intricate properties of the sensory integration such as the arrangements and relative magnitude of excitation and inhibition that are evident in the vast majority of neurons. To study such properties, complementary approaches are needed. STRFs can be converted into a modulation transfer function representation (i.e., using Fourier analysis) that depicts the response strength of the neuron as a function of spectral and temporal sound parameters (Klein et al. 2000; Escabí and Schreiner 2002; Miller et al. 2002; Qiu et al. 2003). The resulting spectral cross section of the modulation transfer function corresponds to the spectral modulation transfer function (sMTF; Fig. 4c). Conversely, the temporal cross section corresponds to the temporal modulation transfer function (tMTF; Fig. 4d). The neuron's best spectral modulation (bSM, also referred to as best ripple density) and best temporal modulation (bTM) correspond to the locations in the modulation transfer function that produce a maximum or optimal response. These high-order response parameters are directly analogous to those obtained with more conventional techniques and are closely linked to the spectral and temporal arrangements of excitation and inhibition (Qiu et al. 2003).

Demonstrating RF transformations with STRFs

White noise analysis and approximations of the different STRF components have been used in both visual and auditory sensory systems to quantify functional transformations along sensory pathways. The contribution of a single thalamic neuron input to a single cortical neuron can be further quantified by using white noise analysis to characterize the STRFs from synaptically paired neurons (Alonso et al. 2001; Miller et al. 2001). With these combined approaches a dramatic transformation from circular symmetric RFs in the retina and thalamus to elongated oriented RFs in the visual cortex has been observed and quantified (Hubel and Wiesel 1962; Hirsch et al. 1998; Ferster and Miller 2000; Alonso et al. 2001). Recently, a similar approach has been taken to study functional transformations along the auditory pathway. First we will summarize some general principles of transformations in sMTFs and tMTFs along the auditory pathway. In the following section, we will describe how quantification of spectrotemporal interdependence can be achieved with white noise analysis techniques.

Spectral modulation transfer functions in mammals

sMTFs between thalamus and cortex have been quantitatively assessed in the auditory pathway with STRF techniques (Depireux et al. 2001; Sen et al. 2001; Miller et al. 2002; Linden et al. 2003; Qiu et al. 2003). Using paired thalamocortical recordings of STRFs, Miller and colleagues found that only 30% of the neurons showed a significant increase in bandwidth (Miller et al. 2001). Thus, the spectral component of the STRF is not systematically elongated in a fashion similar to that seen in the visual pathway. Furthermore, neurons appear to share nearly identical sMTFs at successive stages of processing in the mammalian auditory system. Measurements of the RF bandwidths and bSM in cats shows that these properties are highly conserved between the auditory midbrain (ICC), thalamus (MGBv), and primary auditory cortex (AI) (Miller et al. 2002; Qiu et al. 2003). Within the cortical pathway, Rutkowski and colleagues found significant differences in the spectral component of STRFs across the core and ventrolateral belt areas of the auditory cortex of guinea pigs (Rutkowski et al. 2002). Thus, the cortical pathway appears not to conserve spectral preferences with a direct hierarchical progression.

Temporal modulation transfer functions in mammals

When temporal RF properties are quantified and compared, a general decrease in the temporal following rates is observed within the progression from periphery to central areas (Depireux et al. 2001; Sen et al. 2001; Miller et al. 2002; Linden et al. 2003; Qiu et al. 2003). Auditory midbrain neurons exhibit the fastest temporal dynamics. The ability to follow rapid transient sound modulations is successively reduced in MGBv and AI

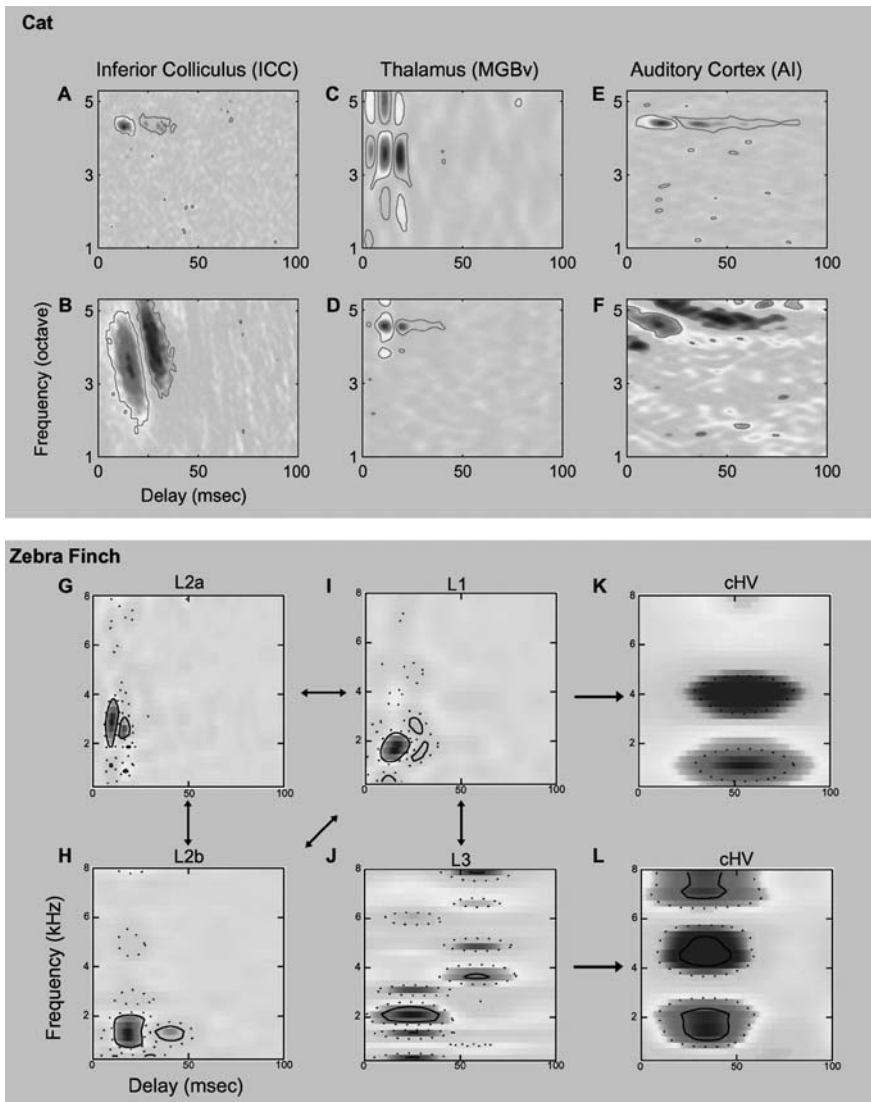


Fig. 5. Differences in RF structure between mutually connected auditory nuclei in the cat and songbird systems. RFs can exhibit simple on-off arrangements (A, D, E, G, H, K) to complex arrangements exhibiting multiple excitatory and inhibitory subregions (C, J, L). Other neurons show selectivity for dynamic frequency transitions and are characterized by obliquely oriented time-frequency excitatory and inhibitory components (B and F). The cat lemniscal (A–F) pathway shows a clear reduction in temporal modulation preferences, although spectral preferences are remarkably similar. In the cat auditory midbrain (ICC, A, and B), neurons exhibit short latency transient RFs with fast modulation preferences, whereas in the primary auditory cortex (AI, E, and F), STRFs are typically of much longer duration. A similar progression from slower to faster and more intricate RF arrangements is observed in the high-order song selective areas of the zebra finch (G–L). *Arrows* indicate the predominant feedforward (*single arrow*) and reciprocal (*bidirectional arrow*) interconnectivity between auditory nuclei in the avian forebrain

(Miller et al. 2002; Qiu et al. 2003). A similar transformation in temporal properties has been demonstrated with more conventional stimuli (Langner and Schreiner 1988; Schreiner and Urbas 1988; Rees and Palmer 1989; Eggermont 1994; Krishna and Semple 2000). A recent study examined differences in cortical STRFs between the thalamic recipient layers of AI and the anterior auditory field (AAF) of mice (Linden et al. 2003). This study showed that the RF differences in these two parallel cortical areas are primarily reflected along the temporal dimension – AAF have faster STRFs. AI and AAF are believed to be hierarchically equal regions, reciprocally connected via layer III and each with one complete cochleotopic representation. This recent finding suggests an important functional difference between the two areas. A reverse hierarchical relationship for temporal preferences has been observed in visual cortical pathways (Heller et al. 1995; Lamme et al. 1998). These studies have extended the classical methods by showing that the timing as opposed to spectral properties account for the dominant receptive field transformation in the ascending auditory pathway.

Spectral and temporal modulation preferences in avian species

A similar set of results was observed in the auditory forebrain circuits of the zebra finch (Fig. 5, G–L). In these studies, Theunissen and colleagues developed a technique for measuring receptive fields with conspecific vocalizations (Theunissen et al. 2000), which they later used to study the transformation between interconnected high-level cortical areas specialized in the audiovocal integration of zebra finch songs (Sen et al. 2001). Their procedure for measuring STRFs with natural sounds is similar to the conventional spike-triggered averaging procedure of Fig. 3 but requires an additional step to remove biases in the spectrotemporal sound structure from bird songs (Fig. 6). This compensation ensures that a broad range of spectrotemporal features is probed so that the STRF reflects the true integration properties of the neuron. Furthermore, it guarantees that there are no systematic distortions in the STRF because of correlations in the natural sound. As for the lemniscal auditory pathway in mammals (Fig. 5, top),

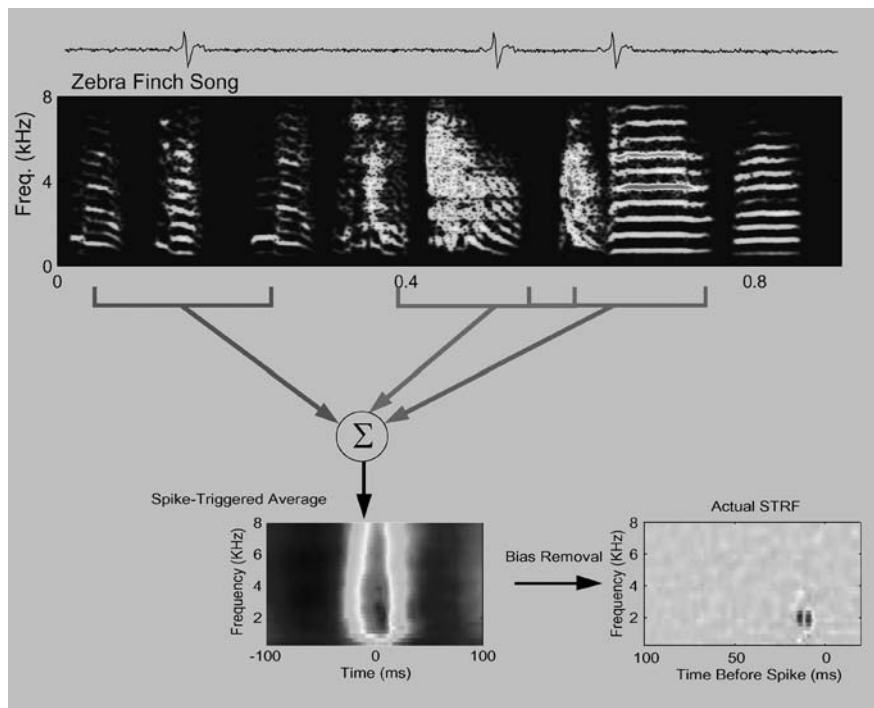


Fig. 6. Illustration of the modified reverse correlation approach for measuring auditory STRFs with natural sounds from a zebra finch neuron. The spike-triggered average waveform obtained with the conventional REVCORR approach of Fig. 3 shows a conspicuous arrangement that extends to times following the action potential. This spike-triggered average waveform reflects correlations in the zebra finch song that distort the structure of the STRF measurement. This measurement, therefore, does not reflect the true RF integration properties of the neuron. To obtain the true STRF, a detrending procedure is applied to the spike-triggered average, which removes the correlations in that natural sound. Following this bias removal, the STRF is constrained to delays before the spike time (at 0 ms), as expected

zebra finch STRFs are fastest at early stages of processing and become successively slower in higher centers (Fig. 5, bottom). RFs in the successive stages of the early mammalian pathway, however, share similar spectral integration properties, whereas RF complexity increases dramatically in successive stages of the avian telencephalon pathway (Sen et al. 2001). Similar results were observed by Rutkowski and colleagues (Rutkowski et al. 2002) in high-level cortical areas. One possible explanation for this finding is that intracortical excitation and inhibition combined with a complex anatomical circuitry in these high-level areas significantly refines spectrotemporal preferences. This, in turn, leads to a dramatic modification of the available STRF arrangements. Such differences in RF organization could potentially play an important role in the analysis of spectrotemporal features in environmental sounds.

Quantifying spectrotemporal interactions

The finding that most of the transformations in the auditory pathway are reflected along the timing properties of the STRF populations suggests that, at the population level, spectral and temporal preferences appear to be independent of each other. This, however, need not be the case for single neurons. As an example, the spatial sensitivity of most visual neurons changes as a function of time, leading to a high selectivity for moving objects. Do auditory neurons exhibit a similar time-dependent frequency tuning, or is frequency tuning largely independent of temporal modulation tuning? If so, how prevalent is this effect? One way to probe this issue is to look at the interdependence between spectral

and temporal dimensions of each STRF (Depireux et al. 2001; Sen et al. 2001; Miller et al. 2002; Linden et al. 2003; Qiu et al. 2003).

One of the major advantages of using white noise analysis techniques is its ability to quantify the interactions or dependencies between spectral and temporal components of a neuron's response to sound. If the STRF can be factorized into separable functions of time and frequency, $STRF(t, f) = H(t) \cdot G(f)$, then spectral integration is independent of temporal processing. Using this approach to measure independence of processing, most studies have found that some neurons do indeed exhibit strong time-frequency inseparability – such as slanted spectrotemporal components (e.g., Fig. 6b and f) – however, this is not the prevailing RF structure in most stations. Instead, many neurons exhibit approximately time-frequency separable structure (Depireux et al. 2001; Sen et al. 2001; Miller et al. 2002; Linden et al. 2003; Qiu et al. 2003), suggesting that temporal and spectral cues are integrated independently of each other. This finding implies that, unlike the overrepresentation of complex RF dynamics that underlies motion sensitivity in the visual system, time-dependent frequency tuning alone is not the predominant mechanism underlying frequency modulation sensitivity in the auditory pathway (Mendelson et al. 1993; Nelken and Versnel 2000). Indeed, a recent study has shown that neuronal sensitivities to time-varying frequency components in primary auditory cortex arise through asymmetric interactions in the timing of excitation and inhibition (Zhang et al. 2003). Thus, sensitivity to time-dependent frequency conjunctions arises through interactions of excitation and inhibition and not through diagonally oriented excitatory components.

5 Limits of the auditory STRF

Linear or nonlinear?

Nonlinear information processing is an essential aspect of all neuronal systems that contributes to their processing efficiency and computing diversity. A recent line of discussion within the auditory community has focused on whether linear modeling approaches, such as the STRF, are potentially well suited for studying various types of nonlinear transformations. It has been suggested that the usefulness of the STRF for studying the auditory system may be significantly limited by its inability to represent nonlinear properties of the STRF. While this is true for certain types of nonlinear transformations, this view is not entirely correct.

The confusion comes from the fact that the STRF is, by definition, a linear model. That is, a prediction of the neuronal response can be obtained by linearly filtering the sound waveform with the STRF. What is not explicitly expressed by this formulation is that the STRF measurement obtained through reverse correlation can actually reflect nonlinear information of the neuron under study (Marmarelis and Marmarelis 1978), so that the STRF structure changes with operating conditions to reflect the neuron's state-dependent changes of the nonlinear transformation.

To illustrate this, consider a cosine signal that is put through various linear and nonlinear transformations (Table 1). When the transformation obeys a linear relationship ($f(x) = x$), the input and output signals are functionally identical, although they may differ by a fixed amplitude-scaling factor. Therefore, the input and output are perfectly correlated, and reverse correlation approaches will accurately reflect this linear transformation. By comparison, for a nonlinear element the output may or may not resemble the input, and this will typically depend on the structure of the nonlinearity. Quadratic nonlinearities, such as for $f(x) = x^2$, distort the input cosine waveform so that it bears no resemblance to the original signal, producing a signal of twice the original frequency. As such, direct reverse correlation methods that rely on linear correlations between the input and output signal will show no evident component in the measured receptive field function, and the technique fails to work. By comparison, cubic-like nonlin-

earities (e.g., $f(x) = x^3$) distort the signal waveform but preserve a significant fraction of the input waveform. This is true even if the operating point is changed, although the correlated component changes with operating point (bottom of Table 1). As such, certain nonlinearities are indeed compatible with reverse correlation and the nonlinear transformation can be systematically evaluated by changing the stimulus operating conditions.

The principles underlying the behavior of the STRF are essentially identical to those described by simple mathematical transformations with simple sinusoid signals. Conceptually, the STRF is compatible with those nonlinear transformations that preserve the precise timing properties between the spectrotemporal stimulus patterns and the neuronal response. The traditional interpretation of these requirements is that the neuron's response should be linearly modulated by the input sound, so that an onset–offset stimulus sequence produces an increase in the neurons firing, followed by a reduction in firing. Although this form of linear rate coding is clearly compatible with the STRF, it is not the only response mode that is amenable to the technique. Precise timing between neuronal events and spectrotemporal sound features could, under various circumstances, play a more important role in the coding of acoustic features. Thus, various types of nonlinearities are structurally compatible with the STRF, assuming that they preserve a precise temporal relationship between the stimulus and the response.

6 Future of the auditory STRF: probing fine temporal and nonlinear response properties

Fine temporal response properties

To illustrate the importance of timing on the STRF construction procedure of Fig. 3, the responses of two auditory midbrain neurons are shown for repeatedly presented segments of a moving ripple stimulus (Fig. 8). The response rastergram of neuron 1 shows a precisely time-locked spike signature down to a few millisecond resolution. For both neurons, the effects of distorting the temporal response pattern are illustrated by adding artificially introduced spike-timing noise to the original

Table 1. Linear and nonlinear transformations (x, x^2, x^3) applied to simple cosine waveforms showing those output (response) components that are compatible with the reverse correlation (REVCORR) approach. Only those terms that are correlated between the input and output show up in the REVCORR mea-

surement (highlighted in *purple*). Changing the operating point by adding a constant DC term to the input, C , for the cubic nonlinearity (x^3) changes the magnitude of the correlated output component. A similar result is also obtained if one changes the operating point of the input by a constant factor for x^2 (not shown)

| Input | Transformation | Output | REVCORR Compatible |
|----------------------------------|----------------|--|--------------------|
| $x = A \cos(\omega \cdot t)$ | $f(x) = x$ | $A \cos(\omega \cdot t)$ | Yes |
| $x = A \cos(\omega \cdot t)$ | $f(x) = x^2$ | $\frac{A^2}{2} + \frac{A^2}{2} \cos(2 \cdot \omega \cdot t)$ | No |
| $x = A \cos(\omega \cdot t)$ | $f(x) = x^3$ | $\frac{3A^3}{4} \cos(\omega \cdot t) + \frac{A^3}{4} \cos(3 \cdot \omega \cdot t)$ | Yes |
| $x = C + A \cos(\omega \cdot t)$ | $f(x) = x^3$ | $C^3 + \frac{3CA^2}{2} + \left[3C^2A + \frac{3A^3}{4} \right] \cos(\omega \cdot t) + \frac{3CA^2}{2} \cos(2 \cdot \omega \cdot t) + \frac{A^3}{4} \cos(3 \cdot \omega \cdot t)$ | Yes |

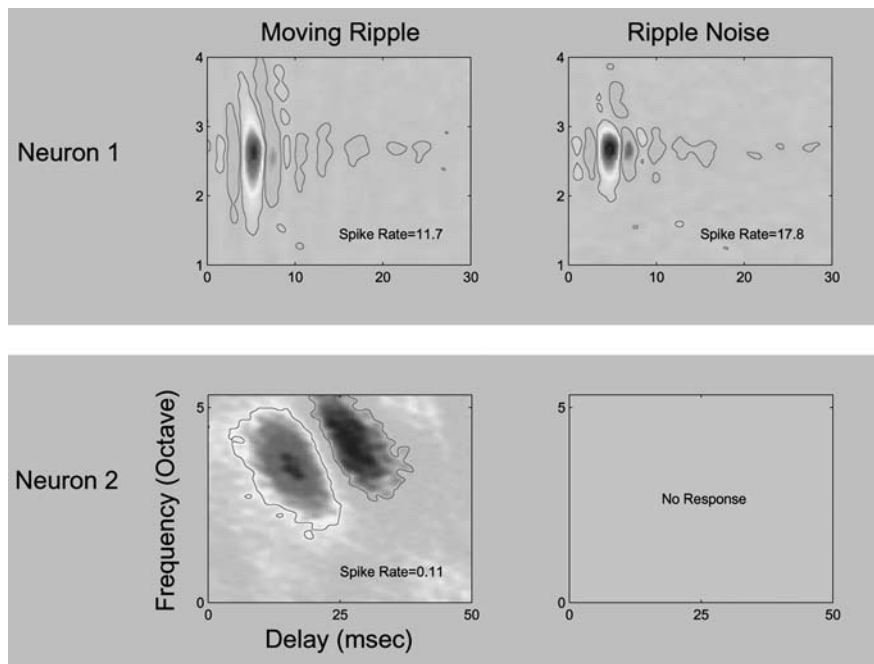


Fig. 7. Timing reliability requirements for STRF construction. The response rastergram showing 100 repeated response trials to dynamic moving ripple for neuron 1 (*top left*, labeled 0 ms) is characterized by phasic responses of a few millisecond resolution. Each spike is represented by a single dot (1-ms resolution). The effects of artificially introducing spike-timing noise on the rastergram are illustrated below the original rastergram (3-, 6-, and 12-ms average errors; shown from *top to bottom*). For this very fast transient neuron, STRF measurements quickly degrade for higher timing errors and are essentially void of structure for errors beyond 3 ms. This shows that precise spike timing is essential for this neuron and these tightly phase-locked neuronal events preserve stimulus-related information down to 3 ms resolution. By comparison, the STRF of neuron 2 (rastergram not shown) tolerates significantly larger spike-timing errors because the spectrotemporal features that it responds to are of much longer duration

response pattern. This intentional introduction of temporal noise has a distinguishable outcome on the STRF structure. In both cases, the RF magnitude becomes weaker and the RF shape is blurred as more spike-timing noise is introduced. This effect, however, is more pronounced for neuron 1 because its tightly synchronized action potentials quickly become misaligned with its preferred sound waveforms and, as a result, the specific temporal relationship between the stimulus modulations and spikes required to generate the STRF is lost. As such, neurons with long RF durations will typically tolerate larger temporal degradation in the spike-timing precision because the stimulus waveforms that the neuron integrates are correlated with the neuronal spikes over a wider range of temporal scales. Clearly, some neurons in the central auditory system do not time-lock to fine temporal modulations (Langner and Schreiner 1988; Lu et al. 2001), and it has been shown that these neurons respond selectively to spectrotemporal modulations, although they do not produce STRFs (Escabi and Schreiner 2002).

Probing nonlinear integration with noise analysis and the STRF

Three recent studies have looked at stimulus-dependent nonlinear integration with complex broadband sounds. By realizing that a linear neuron would produce identical STRFs under a host of operating conditions, these studies have been able to measure state-dependent changes in the RF structure that result from high-order correlations in the stimulus. Theunissen and colleagues (Theunissen et al. 2000) compared the response properties of zebra finch neurons to random tone sequences and zebra finch vocalizations. They found that the RF

structure could be substantially different between these two conditions, reflecting nonlinear inputs that are likely not active for random chord stimuli but become evident with natural sounds. A similar but more pronounced modification of the STRF was also demonstrated in the ICC of cats. In this study, Escabi and Schreiner (2002) compared the response of ICC neurons to dynamic moving ripple (DMR) and ripple noise (RN) sounds (Fig. 3). Although many neurons behaved as nearly linear filters, which produced identical STRFs for the structured DMR and the random RN, a small subset of neurons exhibited a threshold-like suppressive response effect (Fig. 7). These neurons were silenced by the RN sound and responded exclusively to the MR. A third study looked at state-dependent STRF changes in AI that result from the density of spectrotemporal tone pip stimuli (Blake and Merzenich 2002). Using STRF measurements, AI neurons exhibited a sharpening of their frequency selectivity for denser tone pip stimuli. These studies collectively demonstrate how local spectrotemporal sound structure can alter the effective operating regime of the neuron by activating labile nonlinear inputs.

In terms of neuronal mechanisms underlying nonlinear phenomena, there has been little discussion to define the relevant nonlinear processes that are potentially compatible with the STRF approach. Nonlinearities are abundant in the auditory system, ranging from the rectification of hair cell transduction on the basilar membrane to basic biophysical mechanisms of neurons. What sorts of mechanism are compatible with the STRF? The STRF procedure will reflect the composite influence of various neuronal mechanisms assuming that the integration time scale and nonlinear dynamics of each are comparable to the time scales of the relevant stimulus modulation. That is, the nonlinear effects have

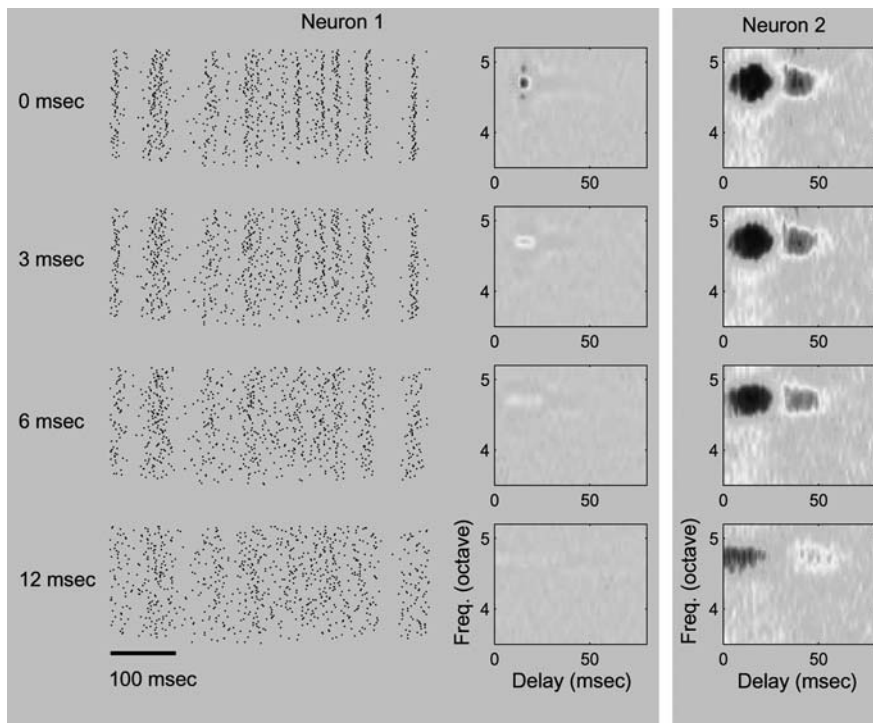


Fig. 8. Linear and nonlinear spectrotemporal integration revealed with STRF analysis. Neurons with approximately linear SRF produce identical STRFs under a variety of stimulus conditions, assuming that equal amounts of energy are presented within its integrating area. Here, neurons were tested with moving ripple and ripple noise stimuli of equal energy in the cat auditory midbrain (Escabi and Schreiner 2002). Neuron 1 exhibits a highly significant STRF with similar magnitudes for both sounds, showing that the neuron is efficiently driven by the sound modulation in both stimuli. Therefore, the neuron behaves approximately as expected for a linear neuron. By comparison, neuron 2 responds only to the structured dynamic ripple sounds and is shut off by the random energy modulations in the ripple noise. This effect shows that high-order properties of the stimulus can significantly modify the operating properties of the neuron

to be sufficiently fast to preserve a one-to-one timing relationship between the stimulus feature and the neuronal response. Such would be the case for the combined effect of fast ionic conductances and the intracellular threshold set by the neuron's reversal potential – assuming that input synaptic activity is of sufficiently large amplitude to produce a suprathreshold response. Other mechanisms such as synaptic suppression, potentiation, and plasticity will likely not show up directly in STRF measurements because these nonlinear effects typically become fully active following a long delay, which would significantly exceed the required timing relationship between stimulus feature and response. However, such long-term changes in neuronal efficacy are analogous to changing the neuron's effective operating regime and could therefore be probed by recording STRFs during a before and after paradigm.

Future studies will need to address the types of biophysical mechanisms and neuronal circuitries underlying specific linear or nonlinear relationships between sound structure and RF organization. Although the simplest way to do this is to identify changes in STRF organization with stimulus operating conditions, there are a number of benefits to be gained by adopting other complementary approaches. In the visual system, for instance, intracellular recordings have been used to parse out the synaptic mechanisms underlying visual feature sensitivities and RF formation in the thalamic input layers of the primary visual cortex (Hirsch et al. 1998; Bringuier et al. 1999; Anderson et al. 2000; Ferster and Miller 2000). These have been complemented by computational neuronal modeling studies aimed at identifying nonlinear mechanisms and the underlying circuitry behind orientation selectivity, contrast invari-

ant tuning, and other nonlinear phenomena in the primary visual cortex. It is likely that similar approaches would be equally fruitful in the auditory system, providing a theoretical framework to develop specific coding hypothesis and paradigms, which could then be tested experimentally.

7 Conclusion

Receptive mapping techniques using white noise analysis and STRF measurements have provided various conceptual advances in our understanding of spectrotemporal feature preferences and the transformation of these across various stages of the ascending auditory system. As auditory circuits in mammals and other species are probed in more detail with complex stimuli, it is likely that the fine details of the different functional transformations will be identified. In particular, STRF approaches will play a role in the analysis of spectrotemporal features of the transformation that could otherwise not be probed directly with simple stimuli.

As for classical stimulus–response characteristics, STRF methods are not sufficient characterizations in themselves, since either of these two approaches fails to fully explain the other. STRFs, however, offer some benefits for measuring neuronal sensitivities to broadband sounds and other complex stimuli including natural sounds. As is evident from recent STRF studies, auditory STRF measurements can be strongly dependent on the spectrotemporal characteristics of probe sound. Future studies will therefore require the development of sounds and methods that can fully exploit the processing complexity of the entire sensory network.

Natural sounds and other complex broadband sounds developed for STRF analysis can partly achieve this by recruiting excitatory and inhibitory inputs across the extent of the sensory epithelium and by activating nonlinear mechanisms that are not fully active with simpler narrowband stimuli.

The auditory STRF is a powerful tool for identifying some of the neuronal dynamics and nonlinear phenomena underlying spectrotemporal integration in the auditory pathway. However, because natural environmental sounds require an extensive number of spectral and temporal parameters to fully describe them, and because auditory neurons exhibit a number of intricate nonlinearities to these stimulus components, it is unlikely that any one tool or approach – including STRF-based methods – will provide the key to understanding the neuronal representation of complex sounds. Ultimately, the biological system itself and the structure in complex environmental sounds constrain the types of analysis techniques that we can use, and each approach will have its own set of advantages and limitations. As such, future endeavors to study the neuronal basis for complex sound analysis will likely require hybrid approaches and sounds to parse out the processing mechanisms and the neuronal circuitry underlying behavior and perception.

Acknowledgements. We thank K. Sen and F.E. Theunissen for generously contributing to the figures. This work was supported by the University of Connecticut Research Foundation and the National Institute of Child Health and Human Development Grant HD-20806.

References

- Alonso JM, Usrey WM et al (2001) Rules of connectivity between geniculate cells and simple cells in cat primary visual cortex. *J Neurosci* 21(11): 4002–4015
- Anderson JS, Lampl I et al (2000) The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science* 290(5498): 1968–1972
- Blake DT, Merzenich MM (2002) Changes of AI receptive fields with sound density. *J Neurophysiol* 88(6): 3409–3420
- Brenowitz EA, Margoliash D et al (1997) An introduction to birdsong and the avian song system. *J Neurobiol* 33(5): 495–500
- Bringuier V, Chavane F, et al (1999) Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* 283(5402): 695–699
- Brosch M, Schreiner CE (1997) Time course of forward masking tuning curves in cat primary auditory cortex. *J Neurophysiol* 77(2): 923–943
- Casseday JH, Ehrlich D et al (1994) Neural tuning for sound duration – role of inhibitory mechanisms in the inferior colliculus. *Science* 264(5160): 847–850
- Cheung SW, Bedenbaugh PH et al (2001) Functional organization of squirrel monkey primary auditory cortex: responses to pure tones. *J Neurophysiol* 85(4): 1732–1749
- De Boer E, de Jongh HR (1967) Correlation studies applied to the frequency resolution of the cochlea. *J Aud Res* 7: 209–217
- De Ruyter van Steveninck RR, Lewen GD et al (1997) Reproducibility and variability in neural spike trains. *Science* 275(5307): 1805–1808
- DeAngelis GC, Ohzawa I et al (1993) Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *J Neurophysiol* 69(4): 1118–1135
- DeCharms RC, Blake DT et al (1998) Optimizing sound features for cortical neurons. *Science* 280(5368): 1439–1443
- Depireux DA, Simon JZ et al (2001) Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol* 85(3): 1220–1234
- Doupe AJ (1997) Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J Neurosci* 17(3): 1147–1167
- Eggermont JJ (1994) Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hear Res* 74(1–2): 51–66
- Escabi MA, Ertel A et al (2003) Intensity dependence of the auditory STRF. *Assoc Res Otolaryngol Abstr* 26: 175
- Escabi MA, Schreiner CE (2002) Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. *J Neurosci* 22(10): 4114–4131
- Ferster D, Miller KD (2000) Neural mechanisms of orientation selectivity in the visual cortex. *Annu Rev Neurosci* 23: 441–471
- Fitzpatrick DC, Kanwal JS et al (1993) Combination-sensitive neurons in the primary auditory cortex of the mustached bat. *J Neurosci* 13(3): 931–940
- Heil P, Rajan R et al (1992) Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. II: Organization of response properties along the 'isofrequency' dimension. *Hear Res* 63(1–2): 135–156
- Heller J, Hertz JA et al (1995) Information flow and temporal coding in primate pattern vision. *J Comput Neurosci* 2(3): 175–193
- Hermes DJ, Aertsen AM et al (1981) Spectro-temporal characteristics of single units in the auditory midbrain of the lightly anaesthetized grass frog (*Rana temporaria L*) investigated with noise stimuli. *Hear Res* 5(2–3): 147–178
- Hirsch JA, Alonso JM et al (1998) Synaptic integration in striate cortical simple cells. *J Neurosci* 18(22): 9517–9528
- Hubel D, Wiesel T (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol Lond* 160: 106–154
- Jones JP, Palmer LA (1987) An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J Neurophysiol* 58(6): 1233–1258
- Klein DJ, Depireux DA et al (2000) Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. *J Comput Neurosci* 9(1): 85–111
- Krishna BS, Semple MN (2000) Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J Neurophysiol* 84(1): 255–273
- Lamme VA, Super H et al (1998) Feedforward, horizontal, and feedback processing in the visual cortex. *Curr Opin Neurobiol* 8(4): 529–535
- Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol* 60(6): 1799–1822
- Liberman MC (1982) The cochlear frequency map for the cat: labeling auditory-nerve fibers of known characteristic frequency. *J Acoust Soc Am* 72(5): 1441–1449
- Linden JF, Liu RC et al (2003) Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J Neurophysiol* (in press)
- Lu T, Liang L et al (2001) Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci* 4(11): 1131–1138
- Marmarelis PZ, Marmarelis VZ (1978) Analysis of physiological systems: the white-noise approach. Plenum, New York
- Marmarelis PZ, Naka KI (1974) Identification of multi-input biological systems. *IEEE Trans Biomed Eng* 21(2): 88–101

- Mendelson JR, Schreiner CE et al (1993) Functional topography of cat primary auditory cortex: responses to frequency-modulated sweeps. *Exp Brain Res* 94(1): 65–87
- Merzenich MM, Knight PL et al (1975) Representation of cochlea within primary auditory cortex in the cat. *J Neurophysiol* 38(2): 231–249
- Merzenich MM, Reid MD (1974) Representation of the cochlea within the inferior colliculus of the cat. *Brain Res* 77(3): 397–415
- Miller LM, Escabi MA et al (2001) Functional convergence of response properties in the auditory thalamocortical system. *Neuron* 32(1): 151–160
- Miller LM, Escabi MA et al (2002) Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. *J Neurophysiol* 87(1): 516–527
- Nelken I, Kim PJ et al (1997) Linear and nonlinear spectral integration in type IV neurons of the dorsal cochlear nucleus. II. Predicting responses with the use of nonlinear models. *J Neurophysiol* 78(2): 800–811
- Nelken I, Versnel H (2000) Responses to linear and logarithmic frequency-modulated sweeps in ferret primary auditory cortex. *Eur J Neurosci* 12(2): 549–562
- Ohlemiller KK, Kanwal JS et al (1996) Facilitative responses to species-specific calls in cortical FM-FM neurons of the mustached bat. *Neuroreport* 7(11): 1749–1755
- Portfors CV, Wenstrup JJ (2001) Topographical distribution of delay-tuned responses in the mustached bat inferior colliculus. *Hear Res* 151(1–2): 95–105
- Qiu A, Schreiner CE et al (2003) Gabor analysis of auditory mid-brain receptive fields: spectro-temporal and binaural composition. *J Neurophysiol* 90(1): 456–476
- Razak KA, Fuzessery ZM et al (1999) Single cortical neurons serve both echolocation and passive sound localization. *J Neurophysiol* 81(3): 1438–1442
- Read HL, Winer JA et al (2001) Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. *Proc Natl Acad Sci USA* 98(14): 8042–8047
- Rees A, Palmer AR (1989) Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. *J Acoust Soc Am* 85(5): 1978–1994
- Rutkowski RG, Shackleton TM et al (2002) Spectrotemporal receptive field properties of single units in the primary, dorso-caudal and ventrorostral auditory cortex of the guinea pig. *Audiol Neurootol* 7(4): 214–227
- Schnupp JW, Msrac-Flogel TD et al (2001) Linear processing of spatial cues in primary auditory cortex. *Nature* 414(6860): 200–204
- Schreiner CE, Langner G (1988) Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J Neurophysiol* 60(6): 1823–1840
- Schreiner CE, Mendelson JR (1990) Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J Neurophysiol* 64(5): 1442–1459
- Schreiner CE, Urbas JV (1988) Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hear Res* 32(1): 49–63
- Sen K, Theunissen FE et al (2001) Feature analysis of natural sounds in the songbird auditory forebrain. *J Neurophysiol* 86(3): 1445–1458
- Suga N, Jen PH (1976) Disproportionate tonotopic representation for processing CF-FM sonar signals in the mustache bat auditory cortex. *Science* 194(4264): 542–544
- Suga N, Simmons JA et al (1975) Peripheral specialization for fine analysis of doppler-shifted echoes in the auditory system of the “CF-FM” bat *Pteronotus parnellii*. *J Exp Biol* 63(1): 161–192
- Theunissen FE, Sen K et al (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 20(6): 2315–2331
- Tusa RJ, Palmer LA et al (1978) The retinotopic organization of area 17 (striate cortex) in the cat. *J Comp Neurol* 177(2): 213–235
- Valentine PA, Eggermont JJ (2003) Onset and steady-state spectrotemporal sensitivity of neurons in cat primary auditory cortex. *Assoc Res Otolaryngol Abstr* 26: 57
- Zhang LI, Tan AY et al (2003) Topography and synaptic shaping of direction selectivity in primary auditory cortex. *Nature* 424(6945): 201–205